Correlations among centrality measures in complex networks

Chang-Yong Lee*

The Department of Industrial Information,

Kongju National University, Chungnam, 340-702 South Korea

(Dated: May 26, 2006)

Abstract

In this paper, we empirically investigate correlations among four centrality measures, originated from the social science, of various complex networks. For each network, we compute the centrality measures, from which the partial correlation as well as the correlation coefficient among measures is estimated. We uncover that the degree and the betweenness centrality are highly correlated; furthermore, the betweenness follows a power-law distribution irrespective of the type of networks. This characteristic is further examined in terms of the conditional probability distribution of the betweenness, given the degree. The conditional distribution also exhibits a power-law behavior independent of the degree which explains partially, if not whole, the origin of the power-law distribution of the betweenness. A similar analysis on the random network reveals that these characteristics are not found in the random network.

PACS numbers: 89.70.+c, 05.45.Df, 87.23.Ge

*Electronic address: clee@kongju.ac.kr

I. INTRODUCTION

The network (or graph) is a useful way of expressing and investigating quantitatively the characteristics of complex systems in various disciplines. It consists of a set of vertices representing entities, and edges associated with connections between vertices [1]. Numerous complex systems can be and have been expressed in terms of networks, and they are often classified by the research field, such as social [2, 3], technological [4, 5], and biological networks [6, 7], to name just a few.

Early researches on the network focused mainly on the regular and the random networks from which many mathematical results for general structural characteristics have been extracted [8]. Recently, due to the availability of computers and the Internet, study on large-scale statistical properties of complex networks has been possible. It was found that many complex networks had distinctive features in common, such as the power-law distribution of the degree and the clique of the network, resulting in the scale-free [9] and the small world networks [10]. These uncovered characteristics, which differ from those of the regular and random networks, was the trigger that brought about considerable advances in the understanding of complex networks, including the development of numerous analysis tools and devising more accurate topological models for the observed networks [1].

More recently, the research on complex networks drifts also toward the community structure of the networks. It has been shown that various complex networks can be organized in terms of the community structure (or modularity), in which groups of vertices that are highly interconnected have looser connections between them. The analysis of these structures has been a topic of intensive investigation in conjunction with many practical relevance, such as finding functional modules in biological networks [11, 12] and identifying communities in the Web in order to construct, for instance, an efficient search engine [13].

Various attempts have been made to find or identify community structures in complex networks [14]. Examples include the hierarchical clustering [15]; methods based on the edge betweenness [16, 17], the edge-clustering via the degree [18], the information centrality [19], and the eigenvector centrality [20]; the information-theoretic approach via the degree [21]. These methods are directly or indirectly related to the centrality measures. Considering that resulting community structure depends on the choice of measures (including the centrality measures) adopted in various schemes, it would be interesting to investigate any relation among the centrality measures.

The centrality (or sociometric status) has been studied particularly in the social science from the perspective of the social connectivity. It is an incarnations of a concept that describes vertices' prominence and/or importance in terms of features of their network environment [22]. It addresses an issue of which individuals are best connected to other or have most influence. This relative importance was quantified by various measures, developed mainly by researchers of the social networks [23]. Different measures for the centrality have been proposed in the social science. Among them, four centrality measures are commonly used in the network analysis: the degree, the closeness, the betweenness, and the eigenvector centrality [22, 23, 24].

In this paper, we empirically investigate correlations among the centrality measures in complex networks to gain some insight into the potential role of the measures in analyzing complex networks. We restrict our analysis to undirected networks, since some of centrality measures, such as the eigenvector centrality, cannot be defined unambiguously for directed networks. We analyze the film actor network, the scientific collaboration network, the neural network of *Caenorhabditis elegans*, the Internet of both the Autonomous System (AS) and the router levels, and protein interaction networks. Analyzed organisms for protein interaction networks are *Saccharomyces cerevisiae*, *Escherichia coli*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Helicobacter pylori*, and *Homo sapiens* [25].

II. CENTRALITY MEASURES

The centrality measures are introduced as a way of specifying and quantifying the centrality concept of a vertex in a network. Furthermore, they are often classified according to the extent to which a vertex has influence on the others: the immediate effects, the mediative effects, and the total effects centrality [22]. Typical examples which belong to each class are: the closeness and degree for the immediate; the betweenness for the mediative; the eigenvector for the total effect centrality. In addition, these measures are argued to be complementary rather than competitive because they stem from the same theoretical foundation [22]. Although the measures are well known, we restate them here for the completeness with the emphasis on their implications.

The degree centrality is the most basic of all measures which counts how many vertices are involved in an interaction. It is defined, for a vertex i, as the number of edges that the vertex has. That is,

$$d_i = \sum_{j=1}^{n} a_{ij} \ , \tag{1}$$

where n is the number of vertices in the network, and $a_{ij} = 1$ if vertices i and j are connected by an edge, $a_{ij} = 0$ otherwise. It measures the opportunity to receive information flowing through the network with everything else being equal. The degree is also a prominent quantity whose distribution follows a power-law distribution in scale-free networks [9].

The eigenvector centrality can be understood as a refined version of the degree centrality in the sense that it recursively takes into account how neighbor vertices are connected. That is, the eigenvector centrality e_i of a vertex i is proportional to the sum of the eigenvector centrality of the vertices it is connected to. It is defined as

$$e_i = \lambda^{-1} \sum_{j} a_{ij} \ e_j \ , \tag{2}$$

where λ is the largest eigenvalue to assure the centrality is non-negative. Thus, e_i is the ith component of the eigenvector associated with the largest eigenvalue λ of the network. While the eigenvector centrality of a network can be calculated via the standard method [26] using the adjacent matrix representation of the network, it can be also computed by an iterative degree calculation, known as the accelerated power method [27]. This method is not only more efficient, but consistent with the spirit of the refined version of the degree centrality.

The closeness centrality stems from the notion that the influence of central vertices spreads more rapidly throughout a network than that of peripheral ones. It is defined, for each vertex i, as

$$c_i = \left(\sum_{j} \ d_{ij} \right)^{-1} \ , \tag{3}$$

where d_{ij} is the length of the shortest path (geodesic) connecting vertices i and j. Thus, the closeness is closely associated with the characteristic path length [10], the average path length of all paths between all pairs of vertices.

The betweenness centrality, or the load [28], is a measure of the influence of a vertex over the flow of information between every pair of vertices under the assumption that information primarily flows over the shortest path between them. It measures the accumulated number of information transmissions that occur through the pass. The removal of high betweenness vertices sometimes results in disconnecting a network. The betweenness centrality of a vertex i is defined as

$$b_i = \sum_{jk}^{n} \frac{g_{jk(i)}}{g_{jk}} \ , \tag{4}$$

where g_{jk} is the number of geodesics between j and k, and $g_{jk(i)}$ is the number of geodesics that pass through i among g_{jk}. Since b_i is of the order $\mathcal{O}(n^2)$, in this paper, we normalize b_i with its maximum value of $(n-1)(n-2)/2$ so that $b_i \in [0,1]$ for all i.

III. CORRELATION ANALYSIS

A. Correlation coefficients and partial correlations

For every network, we compute the four centrality measures so that all four values are assigned to each vertex. The correlation between a pair of different measures can be estimated by the correlation coefficient [29]. More specifically, it is a quantity which measures the linear correlation between vertex-wise pairs of data, $(A,B) = \{(a_i, b_i),\ i = 1, 2, \cdots, n\}$, and is given as

$$R_{AB} = \frac{\sum \left(a_i - \bar{A}\right)\left(b_i - \bar{B}\right)}{n\ \sigma_A\ \sigma_B} \ , \tag{5}$$

where $\bar{A}$ and σ_A are the mean and standard deviation of the measurements of a centrality measure A. The value of R_{AB} ranges from -1 to 1: 1 being totally correlated, and -1 being totally anti-correlated.

Table I shows correlation coefficients estimated between pairs of data obtained from different centrality measures. As shown in Table I, the degree is strongly correlated with the betweenness and less strongly with the eigenvector centrality; whereas the closeness is weakly correlated with the other measures. This implies that the three measures (the degree, the betweenness, the eigenvector centrality) are closely inter-related. In general, correlation coefficients estimated from different variables could be significantly overlapped. That is, a certain amount of correlation found between any two measures may be tied in with correlations with the third.

To take into account this point, we introduce the partial correlation method [30]. The partial correlation is a method that determines the correlation between any two variables under the assumption that each of them is not correlated with the third. That is, it estimates the correlation between two variables while the third variable is held constant. Formally, the partial correlation between variables A and B while holding C constant is given in terms of the corresponding correlation coefficients as

$$R_{AB\cdot C} = \frac{R_{AB} - R_{BC}\ R_{AC}}{\sqrt{(1 - R_{BC}^2)\,(1 - R_{AC}^2)}} \ . \tag{6}$$

We estimate all possible partial correlations for each correlation coefficient, and results are shown in the parentheses of Table I.

From Table I, we find that the partial correlation between the degree and the betweenness, while holding either the eigenvector or the closeness constant, differs little from the correlation coefficient between them. This implies that the strong correlation between the degree and the betweenness is solely due to the two measures by themselves, and little affected by other measures. In contrast, the partial correlation between the betweenness and the eigenvector (or the betweenness and the closeness) while holding the degree constant is anti-correlated. This implies that the positive correlation between the betweenness and the eigenvector (or the betweenness and the closeness) is almost entirely due to correlations with the degree. That is, a *positive* correlation between them would change dramatically to a *negative* correlation if they were not correlated with the degree centrality. Table I also shows that the correlation between the degree and the eigenvector is affected by the betweenness and closeness.

B. Probability distribution of the betweenness

From the correlation analysis, we uncover that the degree and the betweenness are correlated much strongly than other centrality measures. This is, in a sense, expected since vertices of a high degree would have better chance to be included in the shortest path along a pair of vertices. To address the correlation between the degree k and the betweenness b, we relate them, via the Bayes' theorem, as

$$P(b) = \sum_k P(b|k)\ P(k) \ , \tag{7}$$

and focus on the conditional probability distribution $P(b|k)$ of b given k. To obtain reliable statistics for the conditional distribution, we choose the film actor network as an example since it is composed of the largest number of vertices (over 370,000 vertices) in this study. Figure 1(a) shows a few conditional probability distributions $P(b|k)$. As shown in Fig. 1(a), the conditional distribution approximately follows a power-law form with its exponent $f(k)$ depending on k, i.e.,

$$P(b|k) \propto b^{-f(k)} \ . \tag{8}$$

The k-dependent exponent $f(k)$ can be estimated from different degrees k. As Fig. 1(b) suggests, $f(k)$ depends roughly linearly on k. Thus, we have

$$f(k) \approx \alpha k + \beta \ , \tag{9}$$

where parameters α and β can be estimated by the least square fit.

With Eq. (8) and (9), the probability distribution $P(b)$ of the betweenness b can be expressed as

$$P(b) \propto b^{-\beta} \sum_k b^{-\alpha k} P(k) \ . \tag{10}$$

Under the assumption that $P(k)$ does not blow up as k increase, the dominant contribution of the summation comes from small values of k. Thus, to the first approximation, we find that the betweenness follows a power-law distribution, independent of the degree distribution. That is,

$$P(b) \propto b^{-(\alpha+\beta)} \ , \tag{11}$$

with $\alpha + \beta = 2.89$ for the film actor network.

The power-law distribution of the betweenness can also be obtained by the direct estimate of the betweenness distribution. Figure 2 shows betweenness probability distributions of a few networks. Scale-free networks, such as the film actor and the protein interaction network of *D. melanogaster*, have a power-law in the distribution of the betweenness which was first found in Ref. [28]. Considering that the degree is highly correlated with the betweenness, it is not surprising that the betweenness of scale-free networks follows a power-law distribution. From Fig. 2, we also find that the directly estimated exponent 2.36 for the film actor network is close to the derived exponent of $\alpha + \beta = 2.89$.

Moreover, Fig. 2 shows that the power-law distribution of the betweenness is not restricted to the scale-free network, but held true to other types of networks, such as the collaboration network and the neural network of *C. elegans*. Furthermore, as depicted in Fig. 3, the conditional probability distribution of non scale-free networks, for instance, the collaboration network, is also approximately a power-law distribution; furthermore, the exponent of the distribution is insensitive to the degree k.

The power-law of the conditional probability distribution is less clear for networks of small number of vertices. This is probably due to insufficient number of data to obtain reliable statistics. We, however, have seen the power-law of the conditional distribution for networks composed of relatively sufficient number of vertices, irrespective of the type of networks. From this, we may infer that it is the power-law of the conditional probability distribution that is responsible for the power-law nature of the betweenness.

For a comparison, we apply the same analysis as above to the random network. Table II shows correlation coefficients and partial correlations between measures estimated for the random network. In contrast to the real networks, every centrality measure is very strongly correlated with every other measures. This distinctive characteristic, however, changes dramatically once we introduce the partial correlation. From partial correlation estimates, we find that correlation coefficients between all possible pairs of measures, except that between the degree and the betweenness, contain considerable amount of correlation tied in with the other measures. Similar to the real networks, a strong correlation between the degree and the betweenness is nearly maintained when these measures are assumed not to be correlated with the other measures.

We also examine the conditional probability distribution of the betweenness given the degree. A few conditional distributions $P(b|k)$ of the betweenness b given the degree k are depicted in Fig. 4. Unlike complex networks, the distribution is not a power-law, but approximately a Gaussian irrespective of the degree. Since the conditional distribution of the betweenness given the degree does not follow a power-law distribution, we expect that the betweenness distribution of the random network may as well differ from that of real networks. As shown in Fig. 5, it turns out that the betweenness distribution for the random network can be approximated as a log-normal distribution,

$$P(b) = \frac{1}{\sqrt{2\pi}\ \sigma b}\ e^{-(\ln b-\mu)^2/2\sigma^2}\ , \tag{12}$$

where μ and σ are the scale and the shape parameters of the distribution, respectively.

IV. SUMMARY AND CONCLUSION

In this paper, in order to investigate correlations among the measures, we applied four centrality measures (the degree, the closeness, the betweenness, and the eigenvector centrality) to various types of complex networks as well as the random network. We found that the degree was strongly correlated with the betweenness, and the correlation was robust in the sense that the extent of correlation was little affected by the presence of the other measures. This finding was confirmed by estimating the partial correlation between the degree and the betweenness, while holding either the eigenvector or the closeness constant.

Based on the strong correlation existed between the two measures, we further uncovered the characteristics of the betweenness. Not only for scale-free networks but for other types of networks, the conditional distribution of the betweenness given the degree was approximately a power-law which, in turn, played a predominant role in understanding the power-law distribution of the betweenness. This feature was distinct from the random networks in which the conditional distribution was roughly a Gaussian.

Within complex networks, the scale-free network by itself implies the existence of a hierarchy with respect to the degree centrality [11, 31]. Similarly, the power-law distribution of the betweenness may suggest a new potential role of the betweenness in quantifying the hierarchy in conjunction with the community structure [14]. Therefore, it may provide us with feasibility to use the betweenness and/or related quantities as a measure for constructing hierarchical and community structures of complex networks.

Acknowledgments

We like to thank M. Newman for providing us with the scientific collaboration network data. We also appreciate the open sources of various complex network data available at many URLs. This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2005-041-H00052).

[1] For a review of the network theory, see, for example, M. Newman, SIAM Review **45** 167 (2003); R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002); S. Dorogovtsev and J. Mendes, Adv. Phys. **51**, 1079 (2002).

[2] L. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, Proc. Natl. Acad. Sci. USA **97**, 11149 (2000).

[3] M. Newman, Proc. Natl. Acad. Sci. USA **98**, 404 (2001).

[4] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).

[5] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).

[6] N. Guelzim, S. Bottani, P. Bourgine, and F. Kepes, Nature Genetics **31**, 60 (2002).

[7] H. Jeong, B. Tombor, R. Albert, Z. Litvai, and A.-L. Barabási, Nature (London) **407**, 651 (2000).

[8] P. Erdös and A. Rényi, Pulb. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960); B. Bollobás, *Random Graphs* (Academic Press, London, 1985).

[9] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[10] D. Watts and S. Strogatz, Nature (London) **393**, 440 (1998).

[11] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).

[12] A. Rives and T. Galitski, Proc. Natl. Acad. Sci. USA **100**, 1128 (2003).

[13] G. Flake, S. Lawrence, C. Giles, and F. Coetzee, IEEE Computers **35**, 66 (2002).

[14] For a review of the community structure and methods for finding it, see, for example, M.E.J. Newman, Eur. Phys. J. B **38**, 321 (2004).

[15] J. Scott, *Social Network Analysis: A Handbook* (Sage, London, 2000), 2nd ed.; S. Wasserman and K. Faust, *Social Network Analysis* (Cambridge Univ. Press, Cambridge, U.K. 1994).

[16] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. USA, **99**, 7821 (2002).

[17] J. Tyler, D. Wilkinson, and B. Huberman, in *Proceedings of the First International Conference on Communities and Technologies*, edited by M. Huysman, E. Wenger, and V. Wulf (Kluwer, Dordrecht, 2003).

[18] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, Proc. Natl. Acad. Sci. USA **101**, 2658 (2004).

[19] A. Clauset, M.E.J. Newman, and C. Moore, Phys. Rev. E **70**, 066111 (2004).

[20] H. Yang, F. Zhao, W. Wang, T. Zhou, and B. Wang, e-print cond-mat/0508026.

[21] E. Ziv, M. Middendorf, and C. Wiggins, Phys. Rev. E **71**, 046117 (2005).

[22] N. Friedkin, Am. J. Sociol., **96**, 1478 (1991).

[23] L. C. Freeman, Soc. Network. **1**, 215 (1979).

[24] P. Bonacich, J. Math. Sociol., **2**, 113 (1972).

[25] The sources of network data are the following. The film actors was obtained from `http://www.nd.edu/~networks/`; the scientific collaboration data was provided by M. Neuman; the Internet of Autonomous Systems level was obtained from

http://moat.nlanr.net/Routing/rawdata/; the Internet of router level is collected by the Mercator and is available at http://www.isi.edu/scan/mercator/; protein interaction networks data are available at http://dip.doe-mbi.ucla.edu; the somatic nervous system of *Nematode C. elegans* was obtained from http://ims.dse.ibaraki.ac.jp/research/.

[26] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes: The Art of Scientific Computing*, (Cambridge Univ. Press, Cambridge, 2002), chap. 11.

[27] H. Hotelling, Psychometrika **1**, 27 (1936).

[28] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).

[29] See, for example, L. Chao, *Statistics: Methods and Analyses* (McGraw-Hill, New York, 1969).

[30] J. Johnston and J. Dinardo, *Econometric Methods*, (McGraw-Hill, Irwin, 1984), 3rd ed.

[31] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Science **297**, 1551 (2002).

TABLE I: Correlation coefficients and corresponding partial correlations (in the parentheses) between pairs of centrality measures for each network. X stands for the degree centrality; while Y, Z, and W stand for the betweenness, the eigenvector, and the closeness centrality, respectively. Note that the notation for the partial correlation is abbreviated in such a way that corresponding two variables are replaced by a "big dot".

Network	R_{XY} ($R_{\bullet Z}, R_{\bullet W}$)	R_{XZ} ($R_{\bullet Y}, R_{\bullet W}$)	R_{YZ} ($R_{\bullet Y}, R_{\bullet Z}$)	R_{XW} ($R_{\bullet X}, R_{\bullet Z}$)	R_{YW} ($R_{\bullet X}, R_{\bullet Y}$)	R_{ZW} ($R_{\bullet X}, R_{\bullet W}$)
Film actor	0.81 (0.85, 0.81)	0.61 (0.71, 0.59)	0.26 (-0.50, 0.23)	0.31 (0.27, 0.23)	0.20 (-0.10, 0.15)	0.22 (0.04, 0.18)
Internet (AS)	0.98 (0.94, 0.98)	0.82 (0.38, 0.91)	0.79 (-0.12, 0.88)	0.19 (0.16, -0.68)	0.16 (-0.12, -0.65)	0.60 (0.80, 0.79)
Internet (router)	0.58 (0.55, 0.57)	0.36 (0.28, 0.34)	0.23 (0.03, 0.21)	0.29 (0.26, 0.27)	0.15 (-0.03, 0.13)	0.12 (0.02, 0.09)
Collaboration	0.72 (0.71, 0.65)	0.53 (0.52, 0.45)	0.26 (-0.21, 0.14)	0.56 (0.43, 0.49)	0.40 (-0.00, 0.35)	0.33 (0.04, 0.25)
Neural network	0.73 (0.70, 0.59)	0.95 (0.95, 0.74)	0.58 (-0.53, 0.17)	0.90 (0.86, 0.29)	0.58 (-0.26, 0.15)	0.91 (0.37, 0.86)
S. cerevisiae	0.88 (0.83, 0.89)	0.82 (0.74, 0.72)	0.62 (-0.38, 0.57)	0.57 (0.59, 0.02)	0.34 (-0.40, -0.14)	0.68 (0.45, 0.63)
E. coli	0.82 (0.73, 0.82)	0.75 (0.60, 0.86)	0.57 (-0.12, 0.62)	0.20 (0.08, -0.65)	0.18 (0.04, -0.34)	0.68 (0.82, 0.72)
C. elegans	0.96 (0.92, 0.95)	0.74 (0.32, 0.68)	0.71 (-0.03, 0.65)	0.41 (0.22, -0.05)	0.37 (-0.10, -0.09)	0.60 (0.47, 0.51)
D. melanogaster	0.91 (0.74, 0.90)	0.91 (0.72, 0.81)	0.80 (-0.16, 0.72)	0.69 (0.65, 0.15)	0.51 (-0.42, -0.15)	0.71 (0.28, 0.59)
H. pylori	0.94 (0.80, 0.91)	0.86 (0.46, 0.72)	0.82 (0.06, 0.70)	0.68 (0.42, -0.03)	0.60 (-0.15, -0.16)	0.80 (0.57, 0.67)
H. sapiens	0.73 (0.75, 0.69)	0.52 (0.56, 0.52)	0.20 (-0.31, 0.18)	0.37 (0.13, 0.37)	0.39 (0.19, 0.38)	0.10 (-0.12, 0.02)
Average	0.82 (0.78, 0.79)	0.72 (0.57, 0.67)	0.53 (-0.21, 0.46)	0.47 (0.37, 0.04)	0.35 (-0.12, -0.03)	0.52 (0.34, 0.48)

TABLE II: The correlation coefficients and partial correlation between all possible pairs of centrality measures estimated for the random network of different number of vertices, N=1000, 3000, and 6000. For all cases, each vertex has the same average degree $\langle k \rangle = 10$. X stands for the degree centrality; while Y, Z, and W stand for the betweenness, the eigenvector, and the closeness centrality, respectively. The notation for the partial correlation is abbreviated as Table I.

N	R_{XY} $(R_{\bullet Z}, R_{\bullet W})$	R_{XZ} $(R_{\bullet Y}, R_{\bullet W})$	R_{YZ} $(R_{\bullet Y}, R_{\bullet Z})$	R_{XW} $(R_{\bullet X}, R_{\bullet Z})$	R_{YW} $(R_{\bullet X}, R_{\bullet Y})$	R_{ZW} $(R_{\bullet X}, R_{\bullet W})$
1000	0.97 (0.76, 0.86)	0.95 (0.39, 0.61)	0.94 (0.27, 0.69)	0.92 (0.43, -0.07)	0.90 (0.05, -0.25)	0.97 (0.81, 0.86)
3000	0.98 (0.82, 0.93)	0.95 (0.42, 0.56)	0.94 (0.14, 0.72)	0.93 (0.72, 0.21)	0.88 (-0.43, -0.23)	0.96 (0.67, 0.82)
6000	0.98 (0.79, 0.89)	0.95 (0.45, 0.35)	0.94 (0.16, 0.60)	0.95 (0.77, 0.41)	0.90 (-0.43, -0.10)	0.97 (0.69, 0.83)

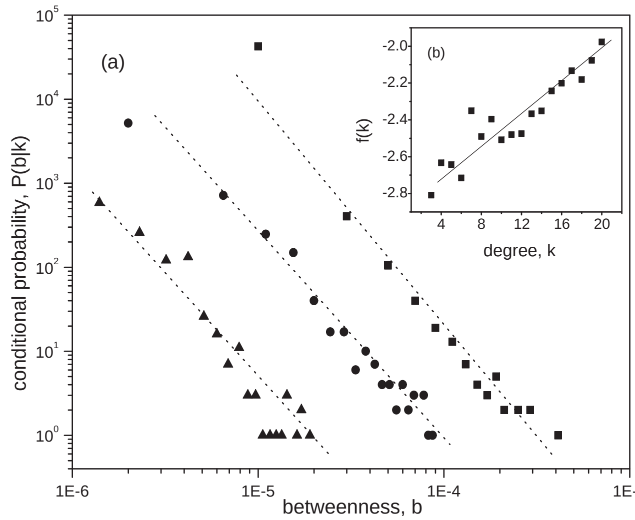

FIG. 1: (a) Log-log plots of the conditional distributions $P(b|k)$ of the betweenness b given the degree k in the film actor network: $k = 3$ (■), $k = 7$ (●), and $k = 10$ (▲). The least-square fits (dotted lines) on the slope of $k = 3$, $k = 7$, and $k = 10$ yield -2.63 ± 0.12, -2.35 ± 0.10, and -2.51 ± 0.18, respectively. Plots for $k = 7$ and 10 are shifted to the left for the display purpose. (b) (inset) The plot of the exponent $f(k)$ in Eq. (8) versus the degree k. Estimated values from the least square fit for Eq. (9) are $\alpha = 0.04 \pm 0.01$ and $\beta = -2.85 \pm 0.05$. The errors associated with the fit are statistical uncertainties based on fitting a straight line.

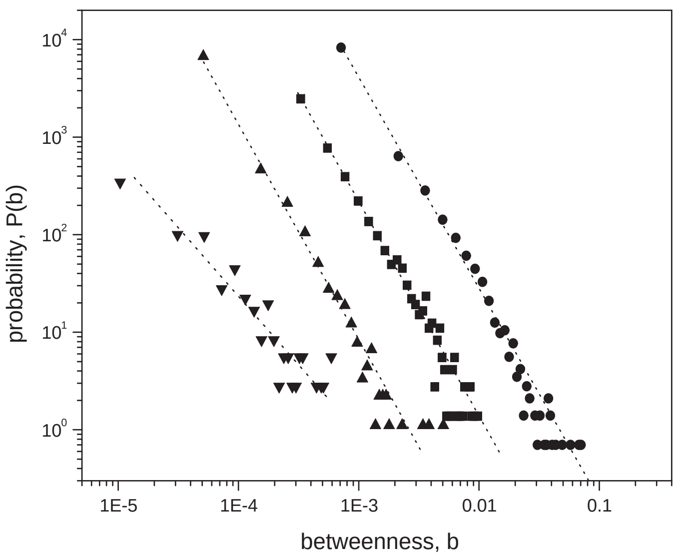

FIG. 2: Log-log plots of betweenness distributions for selected complex networks: the film actor network (■), the collaboration network (•), the protein interaction network of *D. melanogaster* (▲), and the neural network of *C. elegans* (▼). Estimated exponents (dotted lines), by least square fits on slopes, are 2.36 ± 0.10, 2.27 ± 0.08, 2.11 ± 0.12, and 1.31 ± 0.11, respectively. Plots, except for the film actor network, are shifted horizontally for the display purpose.

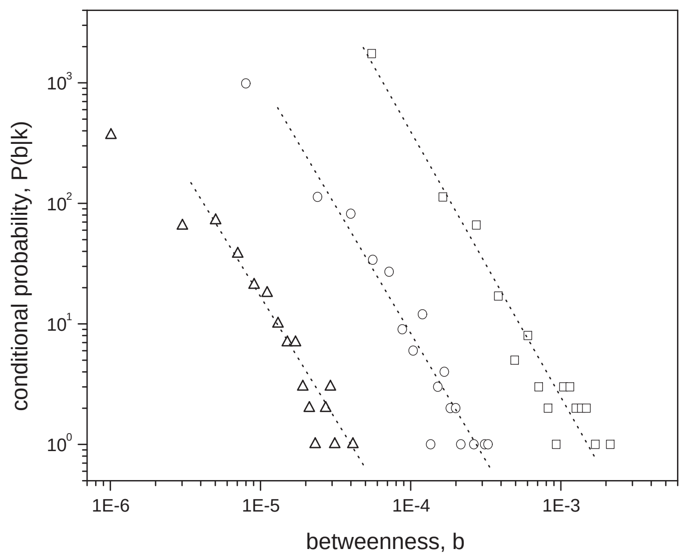

FIG. 3: Log-log plots of the conditional distributions $P(b|k)$ of the betweenness b given degree k of the scientific collaboration network: $k = 4$ (□), $k = 6$ (○), and $k = 9$ (△). The least-square fits (dotted lines) on the slope of the $k = 4$, $k = 6$, and $k = 9$ yield -2.07 ± 0.15, -2.01 ± 0.13, and -2.22 ± 0.19, respectively. Plots for $k = 6$ and 9 are shifted to the left for the display purpose.

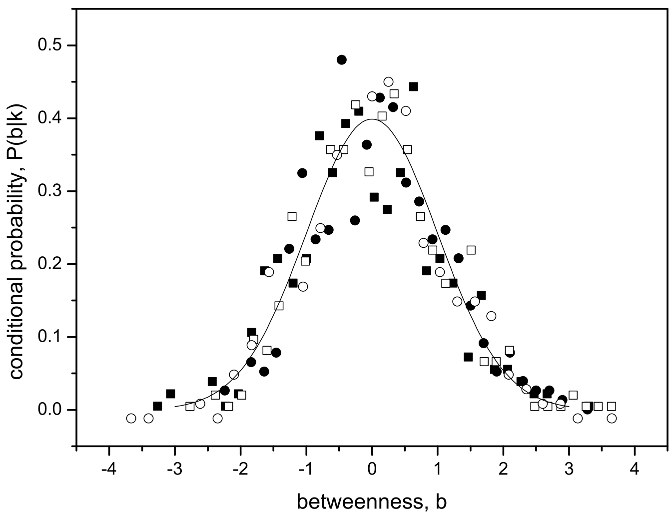

FIG. 4: Plots of the conditional distributions $P(b|k)$ of the betweenness b given the degree k: $k = 8$ (■), $k = 10$ (•), $k = 12$ (□), and $k = 14$ (○) for the random network of 3000 vertices and the average degree $\langle k \rangle = 10$. Each distribution of different k is normalized such that $b \to (b - \bar{b})/\sigma_b$, where $\bar{b}$ and σ_b are the mean and standard deviation of b. By the normalization, all distributions collapse to the standard Gaussian distribution (solid line).

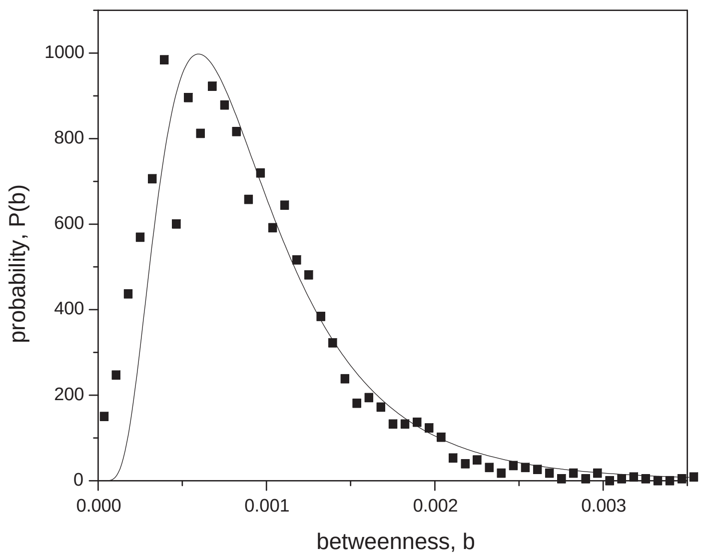

FIG. 5: The distribution of the betweenness for the random network of 3000 vertices and the average degree $\langle k \rangle = 10$, together with the corresponding log-normal fit (solid line). The scale and shape parameters of the log-normal fit are estimated using the maximum likelihood estimate from the data. Estimated the scale and shape parameters are $\hat{\mu} = e^{-0.71}$ and $\hat{\sigma} = 0.57$, respectively.